

- 57 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: (OTHER THAN US): The Council of The Queensland Institute of Medical Research

(US ONLY): HAYWARD Nicholas, SILINS Ginters, GRIMMOND Sean, GARTSIDE Michael and HANCOCK, John

(ii) TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR

(iii) NUMBER OF SEQUENCES: 45

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DAVIES COLLISON CAVE

(B) STREET: 1 LITTLE COLLINS STREET

(C) CITY: MELBOURNE

(D) STATE: VICTORIA

(E) COUNTRY: AUSTRALIA

(F) ZIP: 3000

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT INTERNATIONAL

(B) FILING DATE: 22-MAY-1998

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PO6973

(B) FILING DATE: 23-MAY-1997

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PO6974

(B) FILING DATE: 23-MAY-1997

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PO6972

(B) FILING DATE: 23-MAY-1997

Sub
a2

004450 034500

- 58 -

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PP1459

(B) FILING DATE: 22-JAN-1998

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PP1460

(B) FILING DATE: 22-JAN-1998

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PP1458

(B) FILING DATE: 22-JAN-1998

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUGHES, DR E JOHN L

(C) REFERENCE/DOCKET NUMBER: EJH/AF

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: +61 3 9254 2777

(B) TELEFAX: +61 3 9254 2770

(C) TELEX: AA 31787

003760 "BETHESDA"

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

(2) INFORMATION FOR SEQ ID NO:2:

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AGA AAG GTG ACC AAC CTG TTC TGC TTC GAA CAT CGG GTC AAC GTC TGC 101
Arg Lys Val Thr Asn Leu Phe Cys Phe Glu His Arg Val Asn Val Cys
10 15 20

GAG CAC TGC CTG GTA GCC AAT CAC GCC AAG TGC ATC GTC CAG TCC TAC 149
Glu His Cys Leu Val Ala Asn His Ala Lys Cys Ile Val Gln Ser Tyr
25 30 35 40

CTG CAA TGG CTC CAA GAT AGC GAC TAC AAC CCC AAT TGC CGC CTG TGC 197
Leu Gln Trp Leu Gln Asp Ser Asp Tyr Asn Pro Asn Cys Arg Leu Cys
45 50 55

AAC ATA CCC CTG GCC AGC CGA GAG ACG ACC CGC CTT GTC TGC TAT GAT 245
Asn Ile Pro Leu Ala Ser Arg Glu Thr Thr Arg Leu Val Cys Tyr Asp
60 65 70

CTC TTT CAC TGG GCC TGC CTC AAT GAA CGT GCT GCC CAG CTA CCC CGA 293
Leu Phe His Trp Ala Cys Leu Asn Glu Arg Ala Ala Gln Leu Pro Arg
75 80 85

AAC ACG GCA CCT GCC GGC TAT CAG TGC CCC AGC TGC AAT GGC CCC ATC 341
Asn Thr Ala Pro Ala Gly Tyr Gln Cys Pro Ser Cys Asn Gly Pro Ile
90 95 100

TTC CCC CCA ACC AAC CTG GCT GGC CCC GTG GCC TCC GCA CTG AGA GAG 389
Phe Pro Pro Thr Asn Leu Ala Gly Pro Val Ala Ser Ala Leu Arg Glu
105 110 115 120

- 60 -

AAG CTG GCC ACA GTC AAC TGG GCC CGG GCA GGA CTG GGC CTC CCT CTG Lys Leu Ala Thr Val Asn Trp Ala Arg Ala Gly Leu Gly Leu Pro Leu 125 130 135	437
ATC GAT GAG GTG GTG AGC CCA GAG CCC GAG CCC CTC AAC ACG TCT GAC Ile Asp Glu Val Val Ser Pro Glu Pro Glu Pro Leu Asn Thr Ser Asp 140 145 150	485
TTC TCT GAC TGG TCT AGT TTT AAT GCC AGC AGT ACC CCT GGA CCA GAG Phe Ser Asp Trp Ser Ser Phe Asn Ala Ser Ser Thr Pro Gly Pro Glu 155 160 165	533
GAG GTA GAC AGC GCC TCT GCT GCC CCA GCC TTC TAC AGC CGA GCC CCC Glu Val Asp Ser Ala Ser Ala Ala Pro Ala Phe Tyr Ser Arg Ala Pro 170 175 180	581
CGG CCC CCA GCT TCC CCA GGC CGG CCC GAG CAG CAC ACA GTG ATC CAC Arg Pro Pro Ala Ser Pro Gly Arg Pro Glu Gln His Thr Val Ile His 185 190 195 200	629
ATG GGC AAT CCT GAG CCC TTG ACT CAC GCC CCT AGG AAG GTG TAT GAT Met Gly Asn Pro Glu Pro Leu Thr His Ala Pro Arg Lys Val Tyr Asp 205 210 215	677
ACG CGG GAT GAT GAC CGG ACA CCA GGC CTC CAT GGA GAC TGT GAC GAT Thr Arg Asp Asp Arg Thr Pro Gly Leu His Gly Asp Cys Asp Asp 220 225 230	725
GAC AAG TAC CGA CGT CGG CCG GCC TTG GGT TGG CTG GCC CGG CTG CTA Asp Lys Tyr Arg Arg Arg Pro Ala Leu Gly Trp Leu Ala Arg Leu Leu 235 240 245	773
AGG AGC CGG GCT GGG TCT CGG AAG CGG CCG CTG ACC CTG CTC CAG CGG Arg Ser Arg Ala Gly Ser Arg Lys Arg Pro Leu Thr Leu Leu Gln Arg 250 255 260	821
GCG GGG CTG CTG CTA CTC TTG GGA CTG CTG GGC TTC CTG GCC CTC CTT Ala Gly Leu Leu Leu Leu Gly Leu Leu Phe Leu Ala Leu Leu 265 270 275 280	869
GCC CTC ATG TCT CGC CTA GGC CGG GCC GCA GCT GAC AGC GAT CCC AAC Ala Leu Met Ser Arg Leu Gly Arg Ala Ala Asp Ser Asp Pro Asn 285 290 295	917
CTG GAC CCA CTC ATG AAC CCT CAC ATC CGC GTG GGC CCC TCC TGA Leu Asp Pro Leu Met Asn Pro His Ile Arg Val Gly Pro Ser 300 305 310	962
GCCCCCTTGC TTGTGGCTAG GCCAGCCTAG GATGTGGGTT CTGTGGAGGA GAGGCGGGT	1022
AATGGGGAGG CTGAGGGCAC CTCTTCACTG CCCCTCTCCC TCAAGCCTAA GACACTAAGA	1082
CCCCAGACCC AAAGCCAAGT CCACCAGAGT GGCTCGCAGG CCAGGCCTGG AGTCCCCGTG	1142
GGTCAAGCAT TTGTCTTGAC TTGCTTTCTC CCGGGTCTCC AGCCTCCGAC CCCTCGCCCC	1202
ATGAAGGAGC TGGCAGGTGG AAATAAACAA CAACTTTATT	1242

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 310 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 61 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Gly Leu Cys Lys Cys Pro Lys Arg Lys Val Thr Asn Leu Phe Cys
 1 5 10 15
 Phe Glu His Arg Val Asn Val Cys Glu His Cys Leu Val Ala Asn His
 20 25 30
 Ala Lys Cys Ile Val Gln Ser Tyr Leu Gln Trp Leu Gln Asp Ser Asp
 35 40 45
 Tyr Asn Pro Asn Cys Arg Leu Cys Asn Ile Pro Leu Ala Ser Arg Glu
 50 55 60
 Thr Thr Arg Leu Val Cys Tyr Asp Leu Phe His Trp Ala Cys Leu Asn
 65 70 75 80
 Glu Arg Ala Ala Gln Leu Pro Arg Asn Thr Ala Pro Ala Gly Tyr Gln
 85 90 95
 Cys Pro Ser Cys Asn Gly Pro Ile Phe Pro Pro Thr Asn Leu Ala Gly
 100 105 110
 Pro Val Ala Ser Ala Leu Arg Glu Lys Leu Ala Thr Val Asn Trp Ala
 115 120 125
 Arg Ala Gly Leu Gly Leu Pro Leu Ile Asp Glu Val Val Ser Pro Glu
 130 135 140
 Pro Glu Pro Leu Asn Thr Ser Asp Phe Ser Asp Trp Ser Ser Phe Asn
 145 150 155 160
 Ala Ser Ser Thr Pro Gly Pro Glu Glu Val Asp Ser Ala Ser Ala Ala
 165 170 175
 Pro Ala Phe Tyr Ser Arg Ala Pro Arg Pro Pro Ala Ser Pro Gly Arg
 180 185 190
 Pro Glu Gln His Thr Val Ile His Met Gly Asn Pro Glu Pro Leu Thr
 195 200 205
 His Ala Pro Arg Lys Val Tyr Asp Thr Arg Asp Asp Arg Thr Pro
 210 215 220
 Gly Leu His Gly Asp Cys Asp Asp Asp Lys Tyr Arg Arg Arg Pro Ala
 225 230 235 240
 Leu Gly Trp Leu Ala Arg Leu Leu Arg Ser Arg Ala Gly Ser Arg Lys
 245 250 255
 Arg Pro Leu Thr Leu Leu Gln Arg Ala Gly Leu Leu Leu Leu Gly
 260 265 270
 Leu Leu Gly Phe Leu Ala Leu Leu Ala Leu Met Ser Arg Leu Gly Arg
 275 280 285
 Ala Ala Ala Asp Ser Asp Pro Asn Leu Asp Pro Leu Met Asn Pro His
 290 295 300
 Ile Arg Val Gly Pro Ser
 305 310

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- 62 -

(A) LENGTH: 2415 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 3..2188

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CG ATT TCA TTC CTC GCT CCC CAC AGG TCC CTC TCC CCA AAA TAT TCC	47
Ile Ser Phe Leu Ala Pro His Arg Ser Leu Ser Pro Lys Tyr Ser	
1 5 10 15	
CAT CTT GTC CTA GCC CAT CCC CCA GAC TAT CTC AAG GAC CAG CTG TCC	95
His Leu Val Leu Ala His Pro Pro Asp Tyr Leu Lys Asp Gln Leu Ser	
20 25 30	
CCA CGC CCC CGA CCT CCA CTA GGC CTG TGC CAC CCG CTG CCT GCA GGA	143
Pro Arg Pro Arg Pro Pro Leu Gly Leu Cys His Pro Leu Pro Ala Gly	
35 40 45	
AGA CGC CCG GTC CCG GGC CGG GTT AGC CCC ATG GGA ACG CAG CGC CTG	191
Arg Arg Pro Val Pro Gly Arg Val Ser Pro Met Gly Thr Gln Arg Leu	
50 55 60	
TGT GGC CGC GGG ACT CAA GGC TGG CCT GGC TCA AGT GAA CAG CAC GTC	239
Cys Gly Arg Gly Thr Gln Gly Trp Pro Gly Ser Ser Glu Gln His Val	
65 70 75	
CAG GAG GCG ACC TCG TCC GCG GGT TTG CAT TCT GGG GTG GAC GAG CTG	287
Gln Glu Ala Thr Ser Ser Ala Gly Leu His Ser Gly Val Asp Glu Leu	
80 85 90 95	
GGG GTT CGG TCC GAG CCC GGT GGG AGG CTC CCG GAG CGC AGC CTG GGC	335
Gly Val Arg Ser Glu Pro Gly Gly Arg Leu Pro Glu Arg Ser Leu Gly	
100 105 110	
CCA GCC CAC CCC GCG CCG GCG GCC ATG GCA GGC ACC CTG GAC CTG GAC	383
Pro Ala His Pro Ala Pro Ala Ala Met Ala Gly Thr Leu Asp Leu Asp	
115 120 125	
AAG GGC TGC ACG GTG GAG GAG CTG CTC CGC GGG TGC ATC GAA GCC TTC	431
Lys Gly Cys Thr Val Glu Glu Leu Leu Arg Gly Cys Ile Glu Ala Phe	
130 135 140	
GAT GAC TCC GGG AAG GTG CGG GAC CCG CAG CTG GTG CGC ATG TTC CTC	479
Asp Asp Ser Gly Lys Val Arg Asp Pro Gln Leu Val Arg Met Phe Leu	
145 150 155	
ATG ATG CAC CCC TGG TAC ATC CCC TCC TCT CAG CTG GCG GCC AAG CTG	527
Met Met His Pro Trp Tyr Ile Pro Ser Ser Gln Leu Ala Ala Lys Leu	
160 165 170 175	
CTC CAC ATC TAC CAA CAA TCC CGG AAG GAC AAC TCC AAT TCC CTG CAG	575
Leu His Ile Tyr Gln Gln Ser Arg Lys Asp Asn Ser Asn Ser Leu Gln	
180 185 190	
GTG AAA ACG TGC CAC CTG GTC AGG TAC TGG ATC TCC GCC TTC CCA GCG	623
Val Lys Thr Cys His Leu Val Arg Tyr Trp Ile Ser Ala Phe Pro Ala	
195 200 205	

- 63 -

GAG TTT GAC TTG AAC CCG GAG TTG GCT GAG CAG ATC AAG GAG CTG AAG Glu Phe Asp Leu Asn Pro Glu Leu Ala Glu Gln Ile Lys Glu Leu Lys 210 215 220	671
GCT CTG CTA GAC CAA GAA GGG AAC CGA CGG CAC AGC AGC CTA ATC GAC Ala Leu Leu Asp Gln Glu Gly Asn Arg Arg His Ser Ser Leu Ile Asp 225 230 235	719
ATA GAC AGC GTC CCT ACC TAC AAG TGG AAG CGG CAG GTG ACT CAG CGG Ile Asp Ser Val Pro Thr Tyr Lys Trp Lys Arg Gln Val Thr Gln Arg 240 245 250 255	767
AAC CCT GTG GGA CAG AAA AAG CGC AAG ATG TCC CTG TTG TTT GAC CAC Asn Pro Val Gly Gln Lys Lys Arg Lys Met Ser Leu Leu Phe Asp His 260 265 270	815
CTG GAG CCC ATG GAG CTG GCG GAG CAT CTC ACC TAC TTG GAG TAT CGC Leu Glu Pro Met Glu Leu Ala Glu His Leu Thr Tyr Leu Glu Tyr Arg 275 280 285	863
TCC TTC TGC AAG ATC CTG TTT CAG GAC TAT CAC AGT TTC GTG ACT CAT Ser Phe Cys Lys Ile Leu Phe Gln Asp Tyr His Ser Phe Val Thr His 290 295 300	911
GGC TGC ACT GTG GAC AAC CCC GTC CTG GAG CGG TTC ATC TCC CTC TTC Gly Cys Thr Val Asp Asn Pro Val Leu Glu Arg Phe Ile Ser Leu Phe 305 310 315	959
AAC AGC GTC TCA CAG TGG GTG CAG CTC ATG ATC CTC AGC AAA CCC ACA Asn Ser Val Ser Gln Trp Val Gln Leu Met Ile Leu Ser Lys Pro Thr 320 325 330 335	1007
GCC CCG CAG CGG GCC CTG GTC ATC ACA CAC TTT GTC CAC GTG GCG GAG Ala Pro Gln Arg Ala Leu Val Ile Thr His Phe Val His Val Ala Glu 340 345 350	1055
AAG CTG CTA CAG CTG CAG AAC TTC AAC ACG CTG ATG GCA GTG GTC GGG Lys Leu Leu Gln Leu Gln Asn Phe Asn Thr Leu Met Ala Val Val Gly 355 360 365	1103
GGC CTG AGC CAC AGC TCC ATC TCC CGC CTC AAG GAG ACC CAC AGC CAC Gly Leu Ser His Ser Ser Ile Ser Arg Leu Lys Glu Thr His Ser His 370 375 380	1151
GTT AGC CCT GAG ACC ATC AAG CTC TGG GAG GGT CTC ACG GAA CTA GTG Val Ser Pro Glu Thr Ile Lys Leu Trp Glu Gly Leu Thr Glu Leu Val 385 390 395	1199
ACG GCG ACA GGC AAC TAT GGC AAC TAC CGG CGT CGG CTG GCA GCC TGT Thr Ala Thr Gly Asn Tyr Gly Asn Tyr Arg Arg Arg Leu Ala Ala Cys 400 405 410 415	1247
GTG GGC TTC CGC TTC CCG ATC CTG GGT GTG CAC CTC AAG GAC CTG GTG Val Gly Phe Arg Phe Pro Ile Leu Gly Val His Leu Lys Asp Leu Val 420 425 430	1295
GCC CTG CAG CTG GCA CTG CCT GAC TGG CTG GAC CCA GCC CGG ACC CGG Ala Leu Gln Leu Ala Leu Pro Asp Trp Leu Asp Pro Ala Arg Thr Arg 435 440 445	1343
CTC AAC GGG GCC AAG ATG AAG CAG CTC TTT AGC ATC CTG GAG GAG CTG Leu Asn Gly Ala Lys Met Lys Gln Leu Phe Ser Ile Leu Glu Glu Leu 450 455 460	1391
GCC ATG GTG ACC AGC CTG CGG CCA CCA GTA CAG GCC AAC CCC GAC CTG Ala Met Val Thr Ser Leu Arg Pro Pro Val Gln Ala Asn Pro Asp Leu 465 470 475	1439

- 64 -

CTG AGC CTG CTC ACG GTG TCT CTG GAT CAG TAT CAG ACG GAG GAT GAG Leu Ser Leu Leu Thr Val Ser Leu Asp Gln Tyr Gln Thr Glu Asp Glu 480 485 490 495	1487
CTG TAC CAG CTG TCC CTG CAG CGG GAG CCG CGC TCC AAG TCC TCG CCA Leu Tyr Gln Leu Ser Leu Gln Arg Glu Pro Arg Ser Lys Ser Ser Pro 500 505 510	1535
ACC AGC CCC ACG AGT TGC ACC CCA CCA CCC CGG CCC CCG GTA CTG GAG Thr Ser Pro Thr Ser Cys Thr Pro Pro Pro Arg Pro Pro Val Leu Glu 515 520 525	1583
GAG TGG ACC TCG GCT GCC AAA CCC AAG CTG GAT CAG GCC CTC GTG GTG Glu Trp Thr Ser Ala Ala Lys Pro Lys Leu Asp Gln Ala Leu Val Val 530 535 540	1631
GAG CAC ATC GAG AAG ATG GTG GAG TCT GTG TTC CGG AAC TTT GAC GTC Glu His Ile Glu Lys Met Val Glu Ser Val Phe Arg Asn Phe Asp Val 545 550 555	1679
GAT GGG GAT GGC CAC ATC TCA CAG GAA GAA TTC CAG ATC ATC CGT GGG Asp Gly Asp Gly His Ile Ser Gln Glu Glu Phe Gln Ile Ile Arg Gly 560 565 570 575	1727
AAC TTC CCT TAC CTC AGC GCC TTT GGG GAC CTC GAC CAG AAC CAG GAT Asn Phe Pro Tyr Leu Ser Ala Phe Gly Asp Leu Asp Gln Asn Gln Asp 580 585 590	1775
GGC TGC ATC AGC AGG GAG GAG ATG GTT TCC TAT TTC CTG CGC TCC AGC Gly Cys Ile Ser Arg Glu Glu Met Val Ser Tyr Phe Leu Arg Ser Ser 595 600 605	1823
TCT GTG TTG GGG GGG CGC ATG GGC TTC GTA CAC AAC TTC CAG GAG AGC Ser Val Leu Gly Gly Arg Met Gly Phe Val His Asn Phe Gln Glu Ser 610 615 620	1871
AAC TCC TTG CGC CCC GTC GCC TGC CGC CAC TGC AAA GCC CTG ATC CTG Asn Ser Leu Arg Pro Val Ala Cys Arg His Cys Lys Ala Leu Ile Leu 625 630 635	1919
GGC ATC TAC AAG CAG GGC CTC AAA TGC CGA GCC TGT GGA GTG AAC TGC Gly Ile Tyr Lys Gln Gly Leu Lys Cys Arg Ala Cys Gly Val Asn Cys 640 645 650 655	1967
CAC AAG CAG TGC AAG GAT CGC CTG TCA GTT GAG TGT CGG CGC AGG GCC His Lys Gln Cys Lys Asp Arg Leu Ser Val Glu Cys Arg Arg Arg Ala 660 665 670	2015
CAG AGT GTG AGC CTG GAG GGG TCT GCA CCC TCA CCC TCA CCC ATG CAC Gln Ser Val Ser Leu Glu Gly Ser Ala Pro Ser Pro Ser Pro Met His 675 680 685	2063
AGC CAC CAT CAC CGC GCC TTC AGC TTC TCT CTG CCC CGC CCT GGC AGG Ser His His His Arg Ala Phe Ser Phe Ser Leu Pro Arg Pro Gly Arg 690 695 700	2111
CGA GGC TCC AGG CCT CCA GAG ATC CGT GAG GAG GAG GTA CAG ACG GTG Arg Gly Ser Arg Pro Pro Glu Ile Arg Glu Glu Glu Val Gln Thr Val 705 710 715	2159
GAG GAT GGG GTG TTT GAC ATC CAC TTG TA ATAGATGCTG TGGTTGGATC Glu Asp Gly Val Phe Asp Ile His Leu 720 725	2208
AAGGACTCAT TCCTGCCTTG GAGAAAATAC TTCAACCAGA GCAGGGAGCC TGGGGGTGTC	2268
GGGGCAGGAG GCTGGGGATG GGGGTGGGAT ATGAGGGTGG CATGCAGCTG AGGGCAGGGC	2328

009760 054460

- 65 -

CAGGGCTGGT GTCCTAAGG TTGTACAGAC TCTTGTGAAT ATTTGTATTT TCCAGATGGA 2388
 ATAAAAAGGC CCGTGTAAAT AACCTTC 2415

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 728 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ile Ser Phe Leu Ala Pro His Arg Ser Leu Ser Pro Lys Tyr Ser His
 1 5 10 15
 Leu Val Leu Ala His Pro Pro Asp Tyr Leu Lys Asp Gln Leu Ser Pro
 20 25 30
 Arg Pro Arg Pro Pro Leu Gly Leu Cys His Pro Leu Pro Ala Gly Arg
 35 40 45
 Arg Pro Val Pro Gly Arg Val Ser Pro Met Gly Thr Gln Arg Leu Cys
 50 55 60
 Gly Arg Gly Thr Gln Gly Trp Pro Gly Ser Ser Glu Gln His Val Gln
 65 70 75 80
 Glu Ala Thr Ser Ser Ala Gly Leu His Ser Gly Val Asp Glu Leu Gly
 85 90 95
 Val Arg Ser Glu Pro Gly Gly Arg Leu Pro Glu Arg Ser Leu Gly Pro
 100 105 110
 Ala His Pro Ala Pro Ala Ala Met Ala Gly Thr Leu Asp Leu Asp Lys
 115 120 125
 Gly Cys Thr Val Glu Gly Leu Leu Arg Gly Cys Ile Glu Ala Phe Asp
 130 135 140
 Asp Ser Gly Lys Val Arg Asp Pro Gln Leu Val Arg Met Phe Leu Met
 145 150 155 160
 Met His Pro Trp Tyr Ile Pro Ser Ser Gln Leu Ala Ala Lys Leu Leu
 165 170 175
 His Ile Tyr Gln Gln Ser Arg Lys Asp Asn Ser Asn Ser Leu Gln Val
 180 185 190
 Lys Thr Cys His Leu Val Arg Tyr Trp Ile Ser Ala Phe Pro Ala Glu
 195 200 205
 Phe Asp Leu Asn Pro Glu Leu Ala Glu Gln Ile Lys Glu Leu Lys Ala
 210 215 220
 Leu Leu Asp Gln Glu Gly Asn Arg Arg His Ser Ser Leu Ile Asp Ile
 225 230 235 240
 Asp Ser Val Pro Thr Tyr Lys Trp Lys Arg Gln Val Thr Gln Arg Asn
 245 250 255
 Pro Val Gly Gln Lys Lys Arg Lys Met Ser Leu Leu Phe Asp His Leu
 260 265 270

Publ. No. WO 98/53061

- 66 -

Glu Pro Met Glu Leu Ala Glu His Leu Thr Tyr Leu Glu Tyr Arg Ser
 275 280 285
 Phe Cys Lys Ile Leu Phe Gln Asp Tyr His Ser Phe Val Thr His Gly
 290 295 300
 Cys Thr Val Asp Asn Pro Val Leu Glu Arg Phe Ile Ser Leu Phe Asn
 305 310 315 320
 Ser Val Ser Gln Trp Val Gln Leu Met Ile Leu Ser Lys Pro Thr Ala
 325 330 335
 Pro Gln Arg Ala Leu Val Ile Thr His Phe Val His Val Ala Glu Lys
 340 345 350
 Leu Leu Gln Leu Gln Asn Phe Asn Thr Leu Met Ala Val Val Gly Gly
 355 360 365
 Leu Ser His Ser Ser Ile Ser Arg Leu Lys Glu Thr His Ser His Val
 370 375 380
 Ser Pro Glu Thr Ile Lys Leu Trp Glu Gly Leu Thr Glu Leu Val Thr
 385 390 395 400
 Ala Thr Gly Asn Tyr Gly Asn Tyr Arg Arg Arg Leu Ala Ala Cys Val
 405 410 415
 Gly Phe Arg Phe Pro Ile Leu Gly Val His Leu Lys Asp Leu Val Ala
 420 425 430
 Leu Gln Leu Ala Leu Pro Asp Trp Leu Asp Pro Ala Arg Thr Arg Leu
 435 440 445
 Asn Gly Ala Lys Met Lys Gln Leu Phe Ser Ile Leu Glu Glu Leu Ala
 450 455 460
 Met Val Thr Ser Leu Arg Pro Pro Val Gln Ala Asn Pro Asp Leu Leu
 465 470 475 480
 Ser Leu Leu Thr Val Ser Leu Asp Gln Tyr Gln Thr Glu Asp Glu Leu
 485 490 495
 Tyr Gln Leu Ser Leu Gln Arg Glu Pro Arg Ser Lys Ser Ser Pro Thr
 500 505 510
 Ser Pro Thr Ser Cys Thr Pro Pro Pro Arg Pro Pro Val Leu Glu Glu
 515 520 525
 Trp Thr Ser Ala Ala Lys Pro Lys Leu Asp Gln Ala Leu Val Val Glu
 530 535 540
 His Ile Glu Lys Met Val Glu Ser Val Phe Arg Asn Phe Asp Val Asp
 545 550 555 560
 Gly Asp Gly His Ile Ser Gln Glu Glu Phe Gln Ile Ile Arg Gly Asn
 565 570 575
 Phe Pro Tyr Leu Ser Ala Phe Gly Asp Leu Asp Gln Asn Gln Asp Gly
 580 585 590
 Cys Ile Ser Arg Glu Glu Met Val Ser Tyr Phe Leu Arg Ser Ser Ser
 595 600 605
 Val Leu Gly Gly Arg Met Gly Phe Val His Asn Phe Gln Glu Ser Asn
 610 615 620
 Ser Leu Arg Pro Val Ala Cys Arg His Cys Lys Ala Leu Ile Leu Gly

- 67 -

625 630 635 640
 Ile Tyr Lys Gln Gly Leu Lys Cys Arg Ala Cys Gly Val Asn Cys His
 645 650 655
 Lys Gln Cys Lys Asp Arg Leu Ser Val Glu Cys Arg Arg Arg Ala Gln
 660 665 670
 Ser Val Ser Leu Glu Gly Ser Ala Pro Ser Pro Ser Pro Met His Ser
 675 680 685
 His His His Arg Ala Phe Ser Phe Ser Leu Pro Arg Pro Gly Arg Arg
 690 695 700
 Gly Ser Arg Pro Pro Glu Ile Arg Glu Glu Glu Val Gln Thr Val Glu
 705 710 715 720
 Asp Gly Val Phe Asp Ile His Leu
 725

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2309 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 254..2083

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGATTTCATT CCTCGCTCCC CACAGGTCCC TCTCCCCAAA ATATTCCCAT CTTGTCCTAG 60
 CCCATCCCCC AGACTATCTC AAGGACCAGC TGTCCCCACG CCCCCGACCT CCACTAGGCC 120
 TGTGCCACCC GCTGCCTGCA GGAAGACGCC CGGTCCCGGG CCGGGTTAGC CCCATGGGAA 180
 CGGGGTTCGG TCCGAGCCCG GTGGGAGGCT CCCGAGCGC AGCCTGGGCC CAGCCCACCC 240
 CGCGCCGGCG GCC ATG GCA GGC ACC CTG GAC CTG GAC AAG GGC TGC ACG 289
 1 5 10
 Met Ala Gly Thr Leu Asp Leu Asp Lys Gly Cys Thr
 GTG GAG GAG CTG CTC CGC GGC TGC ATC GAA GCC TTC GAT GAC TCC GGG 337
 Val Glu Glu Leu Leu Arg Gly Cys Ile Glu Ala Phe Asp Asp Ser Gly
 15 20 25
 AAG GTG CGG GAC CCG CAG CTG GTG CGC ATG TTC CTC ATG ATG CAC CCC 385
 Lys Val Arg Asp Pro Gln Leu Val Arg Met Phe Leu Met Met His Pro
 30 35 40
 TGG TAC ATC CCC TCC TCT CAG CTG GCG GCC AAG CTG CTC CAC ATC TAC 433
 Trp Tyr Ile Pro Ser Ser Gln Leu Ala Ala Lys Leu Leu His Ile Tyr
 45 50 55 60
 CAA CAA TCC CGG AAG GAC AAC TCC AAT TCC CTG CAG GTG AAA ACG TGC 481
 Gln Gln Ser Arg Lys Asp Asn Ser Asn Ser Leu Gln Val Lys Thr Cys
 65 70 75
 CAC CTG GTC AGG TAC TGG ATC TCC GCC TTC CCA GCG GAG TTT GAC TTG 529

His	Leu	Val	Arg	Tyr	Trp	Ile	Ser	Ala	Phe	Pro	Ala	Glu	Phe	Asp	Leu		
				80					85					90			
AAC	CCG	GAG	TTG	GCT	GAG	CAG	ATC	AAG	GAG	CTG	AAG	GCT	CTG	CTA	GAC		577
Asn	Pro	Glu	Leu	Ala	Glu	Gln	Ile	Lys	Glu	Leu	Lys	Ala	Leu	Leu	Asp		
		95					100					105					
CAA	GAA	GGG	AAC	CGA	CGG	CAC	AGC	AGC	CTA	ATC	GAC	ATA	GAC	AGC	GTC		625
Gln	Glu	Gly	Asn	Arg	Arg	His	Ser	Ser	Leu	Ile	Asp	Ile	Asp	Ser	Val		
	110					115					120						
CCT	ACC	TAC	AAG	TGG	AAG	CGG	CAG	GTG	ACT	CAG	CGG	AAC	CCT	GTG	GGA		673
Pro	Thr	Tyr	Lys	Trp	Lys	Arg	Gln	Val	Thr	Gln	Arg	Asn	Pro	Val	Gly		
125					130					135					140		
CAG	AAA	AAG	CGC	AAG	ATG	TCC	CTG	TTG	TTT	GAC	CAC	CTG	GAG	CCC	ATG		721
Gln	Lys	Lys	Arg	Lys	Met	Ser	Leu	Leu	Phe	Asp	His	Leu	Glu	Pro	Met		
				145					150					155			
GAG	CTG	GCG	GAG	CAT	CTC	ACC	TAC	TTG	GAG	TAT	CGC	TCC	TTC	TGC	AAG		769
Glu	Leu	Ala	Glu	His	Leu	Thr	Tyr	Leu	Glu	Tyr	Arg	Ser	Phe	Cys	Lys		
			160					165					170				
ATC	CTG	TTT	CAG	GAC	TAT	CAC	AGT	TTC	GTG	ACT	CAT	GGC	TGC	ACT	GTG		817
Ile	Leu	Phe	Gln	Asp	Tyr	His	Ser	Phe	Val	Thr	His	Gly	Cys	Thr	Val		
	175						180					185					
GAC	AAC	CCC	GTC	CTG	GAG	CGG	TTC	ATC	TCC	CTC	TTC	AAC	AGC	GTC	TCA		865
Asp	Asn	Pro	Val	Leu	Glu	Arg	Phe	Ile	Ser	Leu	Phe	Asn	Ser	Val	Ser		
	190					195					200						
CAG	TGG	GTG	CAG	CTC	ATG	ATC	CTC	AGC	AAA	CCC	ACA	GCC	CCG	CAG	CGG		913
Gln	Trp	Val	Gln	Leu	Met	Ile	Leu	Ser	Lys	Pro	Thr	Ala	Pro	Gln	Arg		
205					210					215					220		
GCC	CTG	GTC	ATC	ACA	CAC	TTT	GTC	CAC	GTG	GCG	GAG	AAG	CTG	CTA	CAG		961
Ala	Leu	Val	Ile	Thr	His	Phe	Val	His	Val	Ala	Glu	Lys	Leu	Leu	Gln		
				225					230					235			
CTG	CAG	AAC	TTC	AAC	ACG	CTG	ATG	GCA	GTG	GTC	GGG	GGC	CTG	AGC	CAC		1009
Leu	Gln	Asn	Phe	Asn	Thr	Leu	Met	Ala	Val	Val	Gly	Gly	Leu	Ser	His		
			240					245					250				
AGC	TCC	ATC	TCC	CGC	CTC	AAG	GAG	ACC	CAC	AGC	CAC	GTT	AGC	CCT	GAG		1057
Ser	Ser	Ile	Ser	Arg	Leu	Lys	Glu	Thr	His	Ser	His	Val	Ser	Pro	Glu		
		255					260					265					
ACC	ATC	AAG	CTC	TGG	GAG	GGT	CTC	ACG	GAA	CTA	GTG	ACG	GCG	ACA	GGC		1105
Thr	Ile	Lys	Leu	Trp	Glu	Gly	Leu	Thr	Glu	Leu	Val	Thr	Ala	Thr	Gly		

1. The first step is to identify the problem or question that needs to be addressed. This involves understanding the context and the specific requirements of the task.

AGC CTG CGG CCA CCA GTA CAG GCC AAC CCC GAC CTG CTG AGC CTG CTC Ser Leu Arg Pro Pro Val Gln Ala Asn Pro Asp Leu Leu Ser Leu Leu 350 355 360	1345
ACG GTG TCT CTG GAT CAG TAT CAG ACG GAG GAT GAG CTG TAC CAG CTG Thr Val Ser Leu Asp Gln Tyr Gln Thr Glu Asp Glu Leu Tyr Gln Leu 365 370 375 380	1393
TCC CTG CAG CGG GAG CCG CGC TCC AAG TCC TCG CCA ACC AGC CCC ACG Ser Leu Gln Arg Glu Pro Arg Ser Lys Ser Ser Pro Thr Ser Pro Thr 385 390 395	1441
AGT TGC ACC CCA CCA CCC CGG CCC CCG GTA CTG GAG GAG TGG ACC TCG Ser Cys Thr Pro Pro Pro Arg Pro Pro Val Leu Glu Glu Trp Thr Ser 400 405 410	1489
GCT GCC AAA CCC AAG CTG GAT CAG GCC CTC GTG GTG GAG CAC ATC GAG Ala Ala Lys Pro Lys Leu Asp Gln Ala Leu Val Val Glu His Ile Glu 415 420 425	1537
AAG ATG GTG GAG TCT GTG TTC CGG AAC TTT GAC GTC GAT GGG GAT GGC Lys Met Val Glu Ser Val Phe Arg Asn Phe Asp Val Asp Gly Asp Gly 430 435 440	1585
CAC ATC TCA CAG GAA GAA TTC CAG ATC ATC CGT GGG AAC TTC CCT TAC His Ile Ser Gln Glu Glu Phe Gln Ile Ile Arg Gly Asn Phe Pro Tyr 445 450 455 460	1633
CTC AGC GCC TTT GGG GAC CTC GAC CAG AAC CAG GAT GGC TGC ATC AGC Leu Ser Ala Phe Gly Asp Leu Asp Gln Asn Gln Asp Gly Cys Ile Ser 465 470 475	1681
AGG GAG GAG ATG GTT TCC TAT TTC CTG CGC TCC AGC TCT GTG TTG GGG Arg Glu Glu Met Val Ser Tyr Phe Leu Arg Ser Ser Ser Val Leu Gly 480 485 490	1729
GGG CGC ATG GGC TTC GTA CAC AAC TTC CAG GAG AGC AAC TCC TTG CGC Gly Arg Met Gly Phe Val His Asn Phe Gln Glu Ser Asn Ser Leu Arg 495 500 505	1777
CCC GTC GCC TGC CGC CAC TGC AAA GCC CTG ATC CTG GGC ATC TAC AAG Pro Val Ala Cys Arg His Cys Lys Ala Leu Ile Leu Gly Ile Tyr Lys 510 515 520	1825
CAG GGC CTC AAA TGC CGA GCC TGT GGA GTG AAC TGC CAC AAG CAG TGC Gln Gly Leu Lys Cys Arg Ala Cys Gly Val Asn Cys His Lys Gln Cys 525 530 535 540	1873
AAG GAT CGC CTG TCA GTT GAG TGT CGG CGC AGG GCC CAG AGT GTG AGC Lys Asp Arg Leu Ser Val Glu Cys Arg Arg Arg Ala Gln Ser Val Ser 545 550 555	1921
CTG GAG GGG TCT GCA CCC TCA CCC TCA CCC ATG CAC AGC CAC CAT CAC Leu Glu Gly Ser Ala Pro Ser Pro Ser Pro Met His Ser His His His 560 565 570	1969
CGC GCC TTC AGC TTC TCT CTG CCC CGC CCT GGC AGG CGA GGC TCC AGG Arg Ala Phe Ser Phe Ser Leu Pro Arg Pro Gly Arg Arg Gly Ser Arg 575 580 585	2017
CCT CCA GAG ATC CGT GAG GAG GAG GTA CAG ACG GTG GAG GAT GGG GTG Pro Pro Glu Ile Arg Glu Glu Glu Val Gln Thr Val Glu Asp Gly Val 590 595 600	2065
TTT GAC ATC CAC TTG TAATAGATGC TGTGGTTGGA TCAAGGACTC ATTCTGCCT Phe Asp Ile His Leu 605 610	2120

TGGAGAAAAT ACTTCAACCA GAGCAGGGAG CCTGGGGGTG TCGGGGCAGG AGGCTGGGGA 2180
 TGGGGGTGGG ATATGAGGGT GGCATGCAGC TGAGGGCAGG GCCAGGGCTG GTGTCCCTAA 2240
 GGTGTACAG ACTCTGTGA ATATTGTAT TTTCCAGATG GAATAAAAAG GCCCGTGTA 2300
 TTAACCTTC 2309

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 609 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Ala Gly Thr Leu Asp Leu Asp Lys Gly Cys Thr Val Glu Glu Leu
 1 5 10 15
 Leu Arg Gly Cys Ile Glu Ala Phe Asp Asp Ser Gly Lys Val Arg Asp
 20 25 30
 Pro Gln Leu Val Arg Met Phe Leu Met Met His Pro Trp Tyr Ile Pro
 35 40 45
 Ser Ser Gln Leu Ala Ala Lys Leu Leu His Ile Tyr Gln Gln Ser Arg
 50 55 60
 Lys Asp Asn Ser Asn Ser Leu Gln Val Lys Thr Cys His Leu Val Arg
 65 70 75 80
 Tyr Trp Ile Ser Ala Phe Pro Ala Glu Phe Asp Leu Asn Pro Glu Leu
 85 90 95
 Ala Glu Gln Ile Lys Glu Leu Lys Ala Leu Leu Asp Gln Glu Gly Asn
 100 105 110
 Arg Arg His Ser Ser Leu Ile Asp Ile Asp Ser Val Pro Thr Tyr Lys
 115 120 125
 Trp Lys Arg Gln Val Thr Gln Arg Asn Pro Val Gly Gln Lys Lys Arg
 130 135 140
 Lys Met Ser Leu Leu Phe Asp His Leu Glu Pro Met Glu Leu Ala Glu
 145 150 155 160
 His Leu Thr Tyr Leu Glu Tyr Arg Ser Phe Cys Lys Ile Leu Phe Gln
 165 170 175
 Asp Tyr His Ser Phe Val Thr His Gly Cys Thr Val Asp Asn Pro Val
 180 185 190
 Leu Glu Arg Phe Ile Ser Leu Phe Asn Ser Val Ser Gln Trp Val Gln
 195 200 205
 Leu Met Ile Leu Ser Lys Pro Thr Ala Pro Gln Arg Ala Leu Val Ile
 210 215 220
 Thr His Phe Val His Val Ala Glu Lys Leu Leu Gln Leu Gln Asn Phe
 225 230 235 240
 Asn Thr Leu Met Ala Val Val Gly Gly Leu Ser His Ser Ser Ile Ser
 245 250 255

Sequence of the protein

1

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 832 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

```
(ix) FEATURE:
      (A) NAME/KEY: CDS
      (B) LOCATION: 11..733
```

GCCC	GCC	CGC	ATG	CCG	CCC	TTA	CTG	CCC	CTG	CGC	CTG	TGC	CGG	CTG	TGG	49
			Met	Pro	Pro	Leu	Leu	Pro	Leu	Arg	Leu	Cys	Arg	Leu	Trp	
			1				5					10				
CCC	CGC	AAC	CCT	CCC	TCC	CGG	CTC	CTC	GGA	GCG	GCC	GCC	GGG	CAG	CGG	97
Pro	Arg	Asn	Pro	Pro	Ser	Arg	Leu	Leu	Gly	Ala	Ala	Ala	Gly	Gln	Arg	
	15					20					25					
TCC	AGA	CCC	AGT	ACT	TAT	TAT	GAA	CTG	TTG	GGG	GTG	CAT	CCT	GGT	GCC	145
Ser	Arg	Pro	Ser	Thr	Tyr	Tyr	Glu	Leu	Leu	Gly	Val	His	Pro	Gly	Ala	
30					35					40					45	
AGC	ACT	GAG	GAA	GTT	AAA	CGA	GCT	TTC	TTC	TCC	AAG	TCC	AAA	GAG	CTG	193
Ser	Thr	Glu	Glu	Val	Lys	Arg	Ala	Phe	Phe	Ser	Lys	Ser	Lys	Glu	Leu	
				50					55					60		
CAC	CCA	GAC	CGG	GAC	CCT	GGG	AAC	CCA	AGC	CTG	CAC	AGC	CGC	TTT	GTG	241
His	Pro	Asp	Arg	Asp	Pro	Gly	Asn	Pro	Ser	Leu	His	Ser	Arg	Phe	Val	
			65					70					75			
GAG	CTG	AGC	GAG	GCA	TAC	CGT	GTG	CTC	AGC	CGT	GAG	CAG	AGC	CGC	CGC	289
Glu	Leu	Ser	Glu	Ala	Tyr	Arg	Val	Leu	Ser	Arg	Glu	Gln	Ser	Arg	Arg	
		80					85					90				
AGC	TAT	GAT	GAC	CAG	CTC	CGC	TCA	GGT	AGT	CCC	CCA	AAG	TCT	CCA	CGA	337
Ser	Tyr	Asp	Asp	Gln	Leu	Arg	Ser	Gly	Ser	Pro	Pro	Lys	Ser	Pro	Arg	
	95					100					105					
ACC	ACA	GTC	CAT	GAC	AAG	TCT	GCC	CAC	CAA	ACA	CAC	AGC	TCC	TGG	ACA	385
Thr	Thr	Val	His	Asp	Lys	Ser	Ala	His	Gln	Thr	His	Ser	Ser	Trp	Thr	
110					115					120					125	
CCC	CCC	AAC	GCA	CAG	TAC	TGG	TCC	CAG	TTT	CAC	AGC	GTG	AGG	CCA	CAG	433
Pro	Pro	Asn	Ala	Gln	Tyr	Trp	Ser	Gln	Phe	His	Ser	Val	Arg	Pro	Gln	
				130					135					140		
GGG	CCC	CAG	TTG	AGG	CAG	CAG	CAA	CAC	AAA	CAA	AAC	AAA	CAA	GTG	CTG	481
Gly	Pro	Gln	Leu	Arg	Gln	Gln	Gln	His	Lys	Gln	Asn	Lys	Gln	Val	Leu	
			145				150						155			
GGG	TAC	TGC	CTC	CTC	CTC	ATG	CTG	GCG	GGC	ATG	GGC	CTG	CAC	TAC	ATT	529
Gly	Tyr	Cys	Leu	Leu	Leu	Met	Leu	Ala	Gly	Met	Gly	Leu	His	Tyr	Ile	
		160					165					170				
GCC	TTC	AGG	AAG	GTG	AAG	CAG	ATG	CAC	CTT	AAC	TTC	ATG	GAT	GAA	AAG	577
Ala	Phe	Arg	Lys	Val	Lys	Gln	Met	His	Leu	Asn	Phe	Met	Asp	Glu	Lys	
	175					180					185					

- 73 -

GAT CGG ATC ATC ACA GCC TTC TAC AAC GAA GCC CGG GCA CGG GCC AGG 625
 Asp Arg Ile Ile Thr Ala Phe Tyr Asn Glu Ala Arg Ala Arg Ala Arg
 190 195 200 205

GCC AAC AGA GGC ATC CTT CAG CAG GAG CGA CAA CGG CTA GGG CAG CGG 673
 Ala Asn Arg Gly Ile Leu Gln Gln Glu Arg Gln Arg Leu Gly Gln Arg
 210 215 220

CAG CCG CCA CCA TCC GAG CCA ACC CAA GGC CCC GAG ATC CTG CCC CGG 721
 Gln Pro Pro Pro Ser Glu Pro Thr Gln Gly Pro Glu Ile Val Pro Arg
 225 230 235

GGC GCC GGC CCC TGA GGGGCTC ACCTGGATGG GGCCTGCAGT GCGTTCCCGC 773
 Gly Ala Gly Pro *
 240

TTTGCTTCCT TCCCTGGACG GCCCCTCCC CGAAACGCGC GCAATAAAGT GATTCGCAG 832

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 241 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Pro Pro Leu Leu Pro Leu Arg Leu Cys Arg Leu Trp Pro Arg Asn
 1 5 10 15

Pro Pro Ser Arg Leu Leu Gly Ala Ala Ala Gly Gln Arg Ser Arg Pro
 20 25 30

Ser Thr Tyr Tyr Glu Leu Leu Gly Val His Pro Gly Ala Ser Thr Glu
 35 40 45

Glu Val Lys Arg Ala Phe Phe Ser Lys Ser Lys Glu Leu His Pro Asp
 50 55 60

Arg Asp Pro Gly Asn Pro Ser Leu His Ser Arg Phe Val Glu Leu Ser
 65 70 75 80

Glu Ala Tyr Arg Val Leu Ser Arg Glu Gln Ser Arg Arg Ser Tyr Asp
 85 90 95

Asp Gln Leu Arg Ser Gly Ser Pro Pro Lys Ser Pro Arg Thr Thr Val
 100 105 110

His Asp Lys Ser Ala His Gln Thr His Ser Ser Trp Thr Pro Pro Asn
 115 120 125

Ala Gln Tyr Trp Ser Gln Phe His Ser Val Arg Pro Gln Gly Pro Gln
 130 135 140

Leu Arg Gln Gln Gln His Lys Gln Asn Lys Gln Val Leu Gly Tyr Cys
 145 150 155 160

Leu Leu Leu Met Leu Ala Gly Met Gly Leu His Tyr Ile Ala Phe Arg
 165 170 175

Lys Val Lys Gln Met His Leu Asn Phe Met Asp Glu Lys Asp Arg Ile
 180 185 190

Ile Thr Ala Phe Tyr Asn Glu Ala Arg Ala Arg Ala Arg Ala Asn Arg

- 74 -

195

200

205

Gly Ile Leu Gln Gln Glu Arg Gln Arg Leu Gly Gln Arg Gln Pro Pro
 210 215 220
 Pro Ser Glu Pro Thr Gln Gly Pro Glu Ile Val Pro Arg Gly Ala Gly
 225 230 235 240
 Pro

SEQ ID Nos: 10-18 25-36

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 300 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 170..300

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGATTTTCATT CCTCGCTCCC CACAGGTCCC TCTCCCCAAA ATATTCCCAT CTTGTCCTAG 60
 CCCATCCCCC AGACTATCTC AAGGACCAGC TGTCCCCACG CCCCCGACCT CCACTAGGCC 120
 TGTGCCACCC GCTGCCTGCA GGAAGACGCC CGGTCCCGGG CCGGGTTAG CCC CAT 175
 Pro His
 1
 GGG AAC GGG GTT CGG TCC GAG CCC GGT GGG AGG CTC CCG GAG CGC AGC 223
 Gly Asn Gly Val Arg Ser Glu Pro Gly Gly Arg Leu Pro Glu Arg Ser
 5 10 15
 CTG GGC CCA GCC CAC CCC GCG CCG GCG GCC ATG GCA GGC ACC CTG GAC 271
 Leu Gly Pro Ala His Pro Ala Pro Ala Ala Met Ala Gly Thr Leu Asp
 20 25 30
 CTG GAC AAG GGC TGC ACG GTG GAG GAG CT 300
 Leu Asp Lys Gly Cys Thr Val Glu Glu Leu
 35 40

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

Questions **A**nd **A**nswers

Pro His Gly Asn Gly Val Arg Ser Glu Pro Gly Gly Arg Leu Pro Glu
1 5 10 15

Arg Ser Leu Gly Pro Ala His Pro Ala Pro Ala Ala Met Ala Gly Thr
20 25 30

Leu Asp Leu Asp Lys Gly Cys Thr Val Glu Glu Leu
35 40

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Asp Val Asp Glu Glu Asp Glu Val Glu Asp Ile Glu Phe
1 5 10

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Asp Val Asp Gly Asp Gly His Ile Ser Gln Glu Glu Phe
1 5 10

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids

- 76 -

- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Asp His Asp Arg Asp Gly Phe Ile Ser Gln Glu Glu Phe
 1 5 10

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 13 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Asp Gln Asn Gln Asp Gly Cys Ile Ser Arg Glu Glu Met
 1 5 10

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 13 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Asp Val Asp Met Asp Gly Gln Ile Ser Lys Asp Glu Leu
 1 5 10

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

His Phe Val His Val Ala Glu Lys Leu Leu Gln Leu Gln Asn Phe Asn
 1 5 10 15

Thr Leu Met Ala Val Val Gly Gly Leu Ser His Ser Ser Ile Ser Arg
 20 25 30

(2) INFORMATION FOR SEO ID NO:16:

(ii) MOLECULE TYPE: Peptide

Leu Ala Lys Thr Tyr
35

(ii) MOLECULE TYPE: Peptide

Glu Cys
50

(ii) MOLECULE TYPE: Peptide

His Asn Phe His Glu Thr Thr Phe Leu Thr Pro Thr Thr Cys Asn His

- 78 -

1	5	10	15
Cys	Asn	Lys	Leu
20	Trp	Gly	Ile
	25	Leu	Arg
		Gln	Gly
		Phe	Lys
		30	Cys
			Lys
Asp	Cys	Gly	Leu
35	Ala	Val	His
	40	Ser	Cys
		Cys	Lys
		Ser	Asn
		45	Ala
			Val
			Ala
Glu	Cys		
50			

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 15 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GGGATCCCCC TGGTC

15

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GAATTCGGCA CGAGCCGACG G

21

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 78 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATGGAGCAGA AGCTGATCTC CGAGGAGGAC CTGCCCCGGG CAGCTGGATC CGCAGCCCAC

60

CCCGCGCCGG CGGCCATG

78

(2) INFORMATION FOR SEQ ID NO:22:

- (ii) MOLECULE TYPE: DNA

- Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Pro Gly Ala Ala Gly
1 5 10 15
Ser Ala Ala His Pro Ala Pro Ala Ala Met
20 25

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 33 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

- GGATCCGCAG CCCACCCCGC GCCGGCGGCC ATG

33

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Gly Ser Ala Ala His Pro Ala Pro Ala Ala Met
5 10

- (2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GGACAAAGTG TGTGATGAAC C

21

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CTCATCCTCC GTCTGATACT G

21

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GTAGATGTGG ATCAGCTTGG

20

(2) INFORMATION FOR SEQ ID NO:28: .

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AGGTGGAGAA TGGTCAAGG

19

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GTCATAGTCT GTCTCCTACT

20

- 81 -

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ACATAGACAG CGTGCCTACC

20

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TACAACCTTA GGGACACCAG

20

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TGCTGAGCCT GCTCACGGTG

20

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CAAGTGAACA GCACGTCC

18

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:

009720 "034460

- 82 -

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GACTATCTCA AGGACCAGCT G

21

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GGTTCGGTCC GAGCCCGG

18

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GGAGCGATAC TCCAAGTAGG T

21

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

AGCGGGCCAG GCCCCTTC

18

(2) INFORMATION FOR SEQ ID NO:38:

Sequence of the